EDITORIAL

A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives

Yoav Ben-Shlomo\textsuperscript{a} and Diana Kuh\textsuperscript{b}

What is a Life Course Approach to Chronic Disease Epidemiology?

Over the last few years there has been increasing interest in conceptualizing disease aetiology within a life course framework.\textsuperscript{1,2} This approach is not new to Public Health or unique to epidemiology (see below). However, its current resonance and interest within epidemiology reflects the challenging theoretical framework this approach provides. This issue of the \textit{International Journal of Epidemiology} has several papers with a ‘life course theme’. This accompanying editorial is intended to highlight what we believe are the key conceptual issues around life course epidemiology. We have chosen to use examples from chronic disease epidemiology, but this approach is also applicable within the context of infectious diseases\textsuperscript{3} and wider notions of health and wellbeing.\textsuperscript{4}

We have defined a life course approach to chronic disease epidemiology\textsuperscript{4} as the study of long-term effects on chronic disease risk of physical and social exposures during gestation, childhood, adolescence, young adulthood and later adult life. It includes studies of the biological, behavioural and psychosocial pathways that operate across an individual’s life course, as well as across generations, to influence the development of chronic diseases.

Conceptual Models in Life Course Epidemiology

Conventionally, chronic disease cohort studies recruit subjects in mid-life and follow them up for future disease end-points. The risk of developing disease is then related to baseline exposures or changes in exposure measures ascertained at further follow-ups. Even when baseline measures include early life exposures, such as birthweight and childhood socioeconomic position, these would usually be entered into a multivariable model without much attention to the temporal relationship between variables. Merely the collection of exposure data across the life course is not synonymous with a life course model of disease causation. Surprisingly few epidemiological publications explicitly state the temporal ordering of exposure variables and their inter-relationships, both directly or through intermediary variables, with the outcome measure. One example, where this approach was explicitly undertaken was a study testing the influence of early and later life factors on carotid intima thickness.\textsuperscript{5} This diagramatically ordered classes of variables across the life course. Such an approach is commonplace in structural equation modelling, path analysis and graphical models where prior conceptual representations, before statistical modelling, are standard practice.\textsuperscript{6,7}

Figure 1 illustrates such a conceptualization with respect to adult respiratory disease and/or impaired respiratory function. This figure illustrates many potential pathways between intrauterine growth and adult disease. Such a life course model enables the researcher to explicitly test not only early life course exposures with later disease, but possible pathways with potential intermediaries or confounding factors. Path (a) would represent a predominantly biological pathway whereby impaired fetal development of the lung architecture is associated with future respiratory insults from infectious agents and greater susceptibility to impaired lung function in adulthood and/or chronic obstructive airways disease. Path (b) would be a predominantly social pathway whereby adverse childhood socioeconomic position influences adverse childhood exposures as well as adult socioeconomic position and smoking behaviour. Path (c) reflects a socio-biological pathway whereby adverse childhood socioeconomic position is associated with post-natal lung function and subsequently with poor adult lung function through its effects on immune function and the likelihood of exposure to infectious agents. Path (d) is a bio-social pathway so that repeated childhood infections results in adverse educational attainment and lower adult socioeconomic position. Even such a crude model highlights the complex inter-relationships and rather arbitrary differentiation between biological and social mechanisms. As Krieger\textsuperscript{8} asserts a ‘simplistic division of the social and biological will not suffice’. Even such a simplified model (see Strachan\textsuperscript{9} for detailed discussion) is daunting but importantly challenges both clinical epidemiologists and social scientists to operationalize exposures and conceptualize their inter-relationships across the life course.

As well as integrating biological and psychosocial pathways, a life course approach essentially requires some understanding of the natural history and physiological trajectory of normal biological systems (Figure 2). As can be seen for lung function,
(but it applies to many other continuous physiological measures e.g. muscle strength, cognitive function), different periods across the life course influence phases of biological development, stability or decline. This alternative diagrammatic representation of disease aetiology is important for understanding how exposures may differentially act in critical and/or sensitive periods (see below). For example, an exposure acting in early life may adversely affect lung function during the development period (line B) resulting in a diminished physiological reserve without having any appreciable effect on the rate of decline (line C).

Given the wide range of potential exposures and the importance of the timing of an exposure, it is unsurprising that exposures may affect disease risk in more than one way. We have previously proposed a simple classification (Table 1) of potential life course models of health. Some have misinterpreted a life course approach as precluding a critical period model and only encompassing accumulation of risk. The ‘critical period model’ is when an exposure acting during a specific period has lasting or lifelong effects on the structure or function of organs, tissues and body systems which are not modified in any dramatic way by later experience. This is also known as ‘biological programming’ or it is sometimes referred to as a ‘latency model’ and is the basis of the ‘fetal origins of adult disease’ hypothesis. For example, evidence suggests that poor growth in utero leads to a variety of chronic disorders such as cardiovascular disease, non-insulin dependent diabetes, and hypertension. Exposures acting in later life may still influence disease risk in a simple additive way but it is argued that fetal exposures permanently alter anatomical structures and a variety of metabolic systems.

How ‘Critical’ is a Critical Period?

In its purest form, this model advocates that an exposure in a critical period results in permanent and irreversible damage or disease. A stark example of such an exposure would be maternal exposure to thalidomide in pregnancy and limb development. However, within the context of chronic disease, it is important to distinguish the effects of exposure on structure from those on function. Poor intrauterine development may have an adverse effect on the number of muscle cells that develop in the fetus so that at birth a small growth retarded fetus may have a permanent reduction in the number of muscle cells (structure). Such a child, however, may still compensate by muscle hypertrophy so that in functional terms there may be no evident difference. Such an adaptive facility would make evolutionary sense and may only be unmasked at older ages when the underlying structural deficits may become more important as adaptive processes begin to fail. The same is likely to apply to metabolic and hormonal systems, which may be up or down regulated during fetal life but are still capable of modification by many adult exposures.
The second model extends the ‘critical period’ concept as it recognizes the importance of later life effect modifiers. For example, studies have shown that the relationships of coronary heart disease, high blood pressure and insulin resistance with low birthweight are particularly strong or sometimes only observed for subjects who become obese in childhood, adulthood or both.11,15–17 Such subtle differentiation is important as the critical period may again only be critical for those individuals who experience some other exposure. Whilst statistical interaction is not the same as biological interaction, it is plausible to envisage that a biologically compromised system may only result in pathology with the subsequent addition of other physiological or metabolic stressors.

In contrast, factors that raise disease risk or promote good health may accumulate gradually over the life course, although there may be developmental periods when their effects have greater impact on later health than factors operating at other times (see below). This idea is complementary to the notion of allostatic load18 so that as the number and/or duration of exposures increase, there is increasing cumulative damage to biological systems. Environmental or behavioural insults may cause long-term, gradual damage to health in separate and independent ways, or they may cluster together in socially patterned ways. For example, a subject may experience a variety of independent exposures (such as a road traffic accident, subsequent unemployment and finally the death of a spouse), where each event was unrelated to the proceeding one. However, it is far more common for adverse exposures to be clustered. For example, children living in adverse social circumstances are more likely to be of low birthweight, be exposed to poor diets, experience passive smoke exposure, and have worse educational opportunities. In this case, understanding the effects of childhood social class by identifying specific aspects of the early physical or psychosocial environment (such exposure to air pollution or family conflict) or possible mechanisms (such as nutrition, infection or stress) that are associated with adult disease will provide further aetiological insights. Risk factors at different life stages may also accumulate over time because of ‘chains of risk’ where one adverse (beneficial) exposure or experience tends to lead to another, and so on. For example, unemployment will lead to financial insecurity, which in turn will increase the likelihood of marital conflict and possibly physical abuse leading to marital separation and divorce. These links are probabilistic rather than deterministic but are likely to be sequential. In this scenario it is possible to conceive that each exposure increases the risk of separation in a simply cumulative fashion (‘additive effect’). Alternatively, it may be only the final link in the chain, physical abuse, which results in separation (‘trigger effect’). From a preventive point of view, these chains of events help to identify points of intervention where chains of risk may be broken and a new life course trajectory established. Stopping the additive effect will have health benefits but residual damage will remain. Preventing a trigger effect will more
dramatically abolish any adverse risk associated with exposures experienced earlier in the chain.

Critical Versus Sensitive Periods
The terms critical and sensitive periods are often used loosely in epidemiology without much distinction. In the natural sciences a critical period of development refers to a time window when change in the organization of living systems or subsystems towards increasing complexity, greater adaptivity and more efficient functioning is occurring rapidly and may be most easily modifiable by a variety of factors in a favourable or unfavourable direction.19 The intrinsic changes that occur in critical developmental periods are wholly or partially irreversible; sensitive developmental periods are also times of rapid change but there is more scope to modify or even reverse those changes outside the time window. We suggest the following distinctions for the use of these terms in life course chronic disease epidemiology. In this context, the relevance of changes during a critical period is in respect of their long-term effects on intermediate markers of disease risk or manifestation of disease, usually many years after the critical period. A critical period is defined as a limited time window in which an exposure can have adverse or protective effects on development and subsequent disease outcome. Outside this window, this developmental mechanism for mediating exposure and disease risk is no longer available. A sensitive period is a time period when an exposure has a stronger effect on development and hence disease risk than it would at other times; in other words the same exposure outside this time period may still be associated with increased risk but this association is weaker than during the sensitive period. In epidemiological terms both critical and sensitive periods may be understood as qualitatively different exposure-time interactions. For critical periods, there is no excess disease risk associated with exposure outside this time window, whilst for sensitive periods it is merely weaker. Whilst critical period exposures appear obvious if they act during fetal development, they are not limited in this way. For example, the elevated risk of multiple sclerosis amongst European migrants to South Africa is only observed if migration occurred after the age of 15 years.20 Sensitive period effects for chronic diseases are far harder to demonstrate empirically. Hall et al. discuss how clinical disease severity seen with infectious agents varies with age at exposure.3 Critical periods may be more evident for chronic disease risk associated with developmental mechanisms in biological subsystems whereas sensitive periods are likely to be more common in behavioural development. Most of us are familiar with the ease by which children learn a second language and the difficulties encountered in adulthood.

From Individual to Inter-generational and Population Determinants of Health
Most existing life course studies have limited their scope in examining exposures within a single cohort. Whilst studies in the past have examined disease and/or risk factor concordance rates or correlations across generations, it is only more recently that investigators have started to test inter-generational exposure disease associations. Several studies have now shown that both maternal and paternal cardiovascular mortality are associated with offspring birthweight.21–23 Such associations are complex and may reflect long-term adverse socioeconomic circumstances, maternal health, determined by mother’s own birthweight and childhood growth, lifestyle factors such as smoking, and genomic or epigenetic processes.22

Recently the potential for a life course approach to aid understanding of variations in the health and disease of populations over time, across countries and between social groups has been given more attention.24–27 Davey Smith and his colleagues25 suggest that explanations for social inequalities in cause-specific adult mortality lie in socially patterned exposures at different stages of the life course. A life course perspective is being increasingly used in developing the ‘social inequalities in health’ debate.28–31 More broadly, Keating and Hertzman26 argue that the fundamental processes such as neural sculpting that affect brain and behavioural development interact with the growing chaos in the lives of children and adolescents with long-term effects on human capital and hence the wealth of nations. Leon24 suggests that a life course perspective may help understand the underlying geographical patterns of mortality, particularly East-West differences. He states ‘the assumption that inequalities in health today, whether between or within countries are caused by contemporaneous differences in circumstances of life is not sustainable for a range of important diseases that appear to be driven instead by poor socioeconomic circumstances in early life and childhood’.24

Other Epidemiological Paradigms
The end of the 20th century has seen much debate around the need for epidemiological theory, particularly in relation to health inequalities.32,33 Alongside the development of life course epidemiology with its focus on time, has been a growing interest in eco-social or multi-level models that concurrently view risk factors operating at a variety of hierarchical levels from the macroeconomic to the molecular.33–35 Like life course models, they emphasize the need to have an integrative approach between biological and social factors; the ‘embodiment’ of social phenomena into the biological.8,36 To differing degrees, they also acknowledge the temporal relationship between exposures. The distinction between life course and eco-social approaches has been highlighted by others. Krieger, in her review,8 categorizes a life course approach under the broader banner of psychosocial theory. Whilst a life course approach incorporates psychosocial theories it is far broader than this. Hertzman and colleagues attempt to integrate life course with macro, meso and micro level measures of contemporary society into a ‘unified framework’. We have previously written that a life course approach should attempt to integrate eco-social influences on health.37 In this respect we predict that both models may be increasingly used together. We do not believe that life course poses a ‘rival’ conceptual approach but rather that both temporal and hierarchical approaches are complementary and mutually inclusive. We would also argue that a life course approach is not limited to individuals within a single generation but should intertwine biological and social transmission of risk across generations. It must contextualize any exposure both within a hierarchical structure as well as in relation to geographical and secular differences, which may be unique to that cohort of individuals. We have extended the diagram put
forward by Hertzman and colleagues to illustrate these points (Figure 3). Grandparents, parents and children are linked across generations both by common genetic and/or social influences. The potential role of household, neighbourhood and national influences are illustrated acting across time and across individuals. For example, adverse neighbourhood conditions could affect a mother and her child (A). Similarly national exposures (e.g. war time rationing) may be specific to a single population cohort (B) or period effects may be experienced by all individuals (C).

**Putting Life Course in a Historical and Inter-disciplinary Context**

**Epidemiology**

The idea that childhood is important for adult health is not new in epidemiology or public health but was the prevailing model of health in the first half of the 20th century. The idea that sources of risk to adult health lay in early life was given serious consideration during the years following the Second World War. The childhood origins of adult chronic bronchitis were discussed and investigated using data from maturing longitudinal studies. Investigating the prenatal origins of adult cancer was a natural extension of cancer epidemiology where cohort effects and long-term latent effects of early reproductive or occupational exposures were already recognized in adult cancers, and birth characteristics had been linked to childhood cancers. The investigation of the effects of the famine during the Dutch Hunger Winter of 1944–1945 on subsequent human development, particularly mental performance, remains a classic life course study. Stein and her colleagues succinctly expressed the rationale for life course epidemiology when they wrote that ‘the constitution of each cohort at conception follows from the pattern of fertility at the time, and interacts with the succeeding pattern of prenatal and postnatal experience. The surviving adult population carries

![Diagram of multi-generational schema illustrating the possible influences of hierarchical and life course exposures on disease risk across three related individuals](image)

A = Joint neighbourhood effect of exposure on parent and offspring; B = Childhood cohort effect on grandparent; C = Period effect influencing all three generations

**Figure 3** Multi-generational schema illustrating the possible influences of hierarchical and life course exposures on disease risk across three related individuals
the imprint of these favourable and unfavourable experiences during development; morbidity has marked them, and mortality has thinned their ranks, in a way specific to each cohort. Similarly, Jerry Morris in his classic *Uses of Epidemiology* also considered the importance of inter-generational influences, programming and childhood precursors of adult disease (see discussion by Davey Smith). In cardiovascular epidemiology, the adult lifestyle model dominated the post-war era (with some notable exceptions) but interest in early life factors (such as hypertension, smoking, fatty diets and obesity) in cohorts of children. This was then followed by research that linked poor childhood living conditions (see accompanying article by Forsdahl) and accompanying commentaries in this issue and impaired early development to adult cardiovascular disease.

**Other scientific disciplines**

The most pervasive Western model of human development has been the idea that ‘the first few years of life necessarily have crucial effects upon later development and adult characteristics’. The dominant theories of development at the beginning of the 20th century saw development as an invariant sequence of developmental stages, where change was cumulative and usually irreversible. The role of the environment was to provide the initial stimuli and appropriate opportunities for growth and maturation to unfold. Environmental insults or lack of appropriate stimuli during critical periods at the embryonic stage or during infancy, or during childhood more generally could have long-term effects on physical or intellectual development. Once maturity was reached, systems were less able to be manipulated.

From the late 1960s, this early life paradigm was increasingly questioned. It was suggested that continuity of the environment, rather than early experiences per se lay behind evidence of long-term effects on adult characteristics. Nearly every demonstration of a critical period in behavioural development (whether avian imprinting, social behaviour or language acquisition) was followed by a demonstration of some behavioural recovery from the effects of critical period exposure or deprivation. Evidence for irreversible biological changes during critical periods of growth was stronger. Partly as a response to this controversy, a life span perspective in developmental psychology emerged in the 1970s, where psychological development was seen as a lifelong process. The consequences of early developmental experiences could be transformed again and again by later experiences, and the course of development remained malleable into old age. In the UK, Rutter argued that ‘simplistic concepts of immutable effects need to be put aside and replaced by more dynamic notions of the continuing interplay over time ..’ but acknowledged substantial continuities between early experiences and adult psychological functioning. Chain reactions (what we have called ‘chains of risk’) provide an explanation for these continuities, whereby one ‘bad’ thing leads to another, or, conversely, a good experience makes it more likely that another one will be encountered.

Demography and sociologists have investigated how the individual’s life course is shaped by institutions and culture, historical and social change, and changes in individual ageing processes (such as the increase in life expectancy). At the population level cohort effects occur because people of different ages and those who occupy different roles are differentially exposed to and influenced by particular social and economic changes, such as the economic depression of the 1930s. The life course perspective in sociology has encouraged life span psychologists to consider how the individual life course is embedded in the sociohistorical and biocultural context. Changing individuals must be studied in a changing world.

The emerging life course perspective in human biology linking development to ageing, has focused on the ways in which early environmental factors can have influences on human form and function across the life span. This approach has been fostered by biological anthropologists and epidemiologists with their traditional focus on variability, its extent, causes and functional or disease consequences, and by those searching for lifelong precursors of ageing. Selection, optimization and adaptation in the face of adversity and subsequent resilience and reproductive success, but at the expense of later disease.

Thus, there is a different emphasis in interdisciplinary developmental science today compared with the pre-war era. There has been a shift from homogeneity, continuity and universality of developmental processes to heterogeneity, discontinuity and context specific development. The search is for the range of plasticity and its age-associated changes and constraints. This research holds the promise of linking epidemiological observations to underlying mechanisms.

**Empirical Challenges in Undertaking Life Course Analyses**

Adopting a life course approach presents major challenges for both the design and analysis of epidemiological studies. Few researchers are fortunate enough to have access to a birth cohort study with repeat measures of both psychosocial and biological exposures. Such studies can examine common disease outcomes or continuous measures of function, but may be restricted with less common diseases. Frequently such studies will have either no data, especially for biological mechanisms, or missing data in subsets of participants. They clearly remain one of the most powerful ways to test life course models. In addition, cohorts from the same population but different time periods can directly measure secular trends and can provide invaluable insights into whether exposure-disease associations differ across time.

Historical cohort and record linkage studies provide an efficient method to test early life exposures. However, they are limited to usually one or two exposures acting in a very specific time window and frequently have limited or no data on other periods of the life course. This means that possible periods of aetiological importance may be identified but interactions between early and later life exposures or accumulation models cannot be tested. Case-control studies are also possible and can either use proxy measures for early life exposures or retrospectively recalled exposures. Much work remains to be done to quantify the validity of early life exposures reported in adulthood. Occasionally, natural experiments, such as war-time malnutrition, can be used to test extreme time-limited exposure acting during specific time windows. Similarly, randomized controlled trials, usually designed for different purposes, remain
powerful tools to test specific mechanisms or pathways both for biological exposures or social interventions.96–98

Whilst obtaining adequate exposure measures across the life course may be problematic, so is the analysis of such large and complex datasets. We anticipate that as interest in life course methods increases, newer techniques will evolve to help us empirically test some of the theoretical models proposed above. Similarly, techniques already in existence but perhaps currently under-utilized in conventional epidemiological analyses, for example structural equation modelling, path analysis, G-estimation and multi-level modelling, will become more widespread.

Conclusions

Life course epidemiology has challenged the complacency of the adult lifestyle model of chronic disease risk. It has particularly acted as a catalyst for new research in the area of social inequalities in health and has helped bridge biological, psychological and social models of disease causation. A life course approach is paradoxical as on the one hand it is intuitively obvious (do we really need research to demonstrate risk accumulation?), and yet on the other hand is empirically complex (do we actually have much evidence in support of these models?). It remains to be seen whether as researchers we can cope with all this complexity. The future value of a life course approach will depend for its success on elucidating new mechanisms and disease pathways as well as its ability to explain social, geographical and temporal patterns of disease distribution.

Acknowledgements

The authors would like to thank the organizers of the Pisa conference in 1999 where some of these ideas were first presented and Hilary Graham, Johan Halleqvist, Clyde Hertzman, John Lynch, Chris Power for useful discussions at the ESF Summer School Life course workshop, Coimbra, Portugal 2001.

References


285_C01-121  8/4/2002 3:34 pm  Page 292    (Black plate)

292 INTERNATIONAL JOURNAL OF EPIDEMIOLOGY

43 MacMahon B. Breast cancer at menopausal ages: an explanation of observed incidence changes. Cancer 1957;10:1037–44.


